Molecular characterisation of the *rec*A locus in clinical isolates of verocytotoxigenic *E. coli* O157:H7

J. E. MOORE*, M. WATABE*, B. C. MILLAR*, P. J. ROONEY*, A. LOUGHREY*, C. E. GOLDSMITH*, M. A. S. McMAHON†, D. A. McDOWELL† and P. G. MURPHY*,‡

*Northern Ireland Public Health Laboratory, Department of Bacteriology, Belfast City Hospital, Lisburn Road, Belfast; †School of Biomedical Sciences, University of Ulster, Jordanstown, Shore Road, Newtownabbey, Co. Antrim, Northern Ireland; and ‡Department of Medical Microbiology, The Adelaide & Meath Hospital, Tallaght, Dublin 24, Ireland.

Accepted: 19 June 2008

ABSTRACT

Molecular epidemiology of verocytoxigenic *Escherichia coli* O157:H7 is important to help elucidate reservoirs and modes of transmission, particularly between animals and humans. As the *rec*A gene locus is now beginning to gain application in bacterial genotyping schemes, and as it has not been examined previously in *E. coli* O157 isolates, this study aims to examine potential polymorphic variation as a possible epidemiological marker for the subspecies characterisation of clinically significant verocytotoxigenic *E. coli* O157:H7. A novel polymerase chain reaction (PCR) assay was designed to target a 638 bp region of the *rec*A gene in *E. coli* O157 isolates. The PCR amplification of genomic DNA from extracted organisms was able to generate an amplicon of the expected size (approximately 638 bp) for all *E. coli* O157:H7 examined (n=80), as well as for other non-O157 *E. coli* and other members of the Enterobacteriaeceae including *Citrobacter, Hafnia, Shigella, Enterobacter* and *Providencia*. Subsequent restriction fragment length polymorphism (RFLP) and single-stranded conformational polymorphism (SSCP) analyses of these *rec*A amplicons were able to differentiate *E. coli* O157 from the organisms examined, but were unable to distinguish between 79 isolates of wild-type *E. coli* O157, suggesting a highly conserved *rec*A gene structure within the local population of organisms examined.

KEY WORDS: Escherichia coli O157-H7.
Genes.
Polymerase chain reaction.
Polymorphism, restriction fragment length.
Polymorphism, single-stranded conformational

Introduction

Since verocytotoxin-producing *Escherichia coli* (VTEC) was first recognised as an important gastrointestinal bacterial pathogen in humans, mainly associated with zoonotic acquisition from animals, its microbiological characteristics, pathogenicity and epidemiology have been studied worldwide. The severity of symptoms with *E. coli* O157 infection may vary from mild diarrhoea to haemolytic uraemic syndrome (HUS). Following large outbreaks in the USA,[1] Japan[2] and Scotland,[3] *E. coli* O157:H7, in particular, has been monitored closely.

In Northern Ireland, reports of infection with *E. coli* O157 have risen from a few cases in the early 1990s (n=1 [1992], n=2 [1993]) to 54 reports in 1999, which remains the highest yearly number of reports to date, where reports approximately doubled from 1997 to 1999.

Several methods have been used to subtype bacterial isolates and include serotyping,[4] biotyping,[5] phage typing[6] and molecular biological techniques such as the random amplification of polymorphic DNA polymerase chain reaction (RAPD-PCR).[7] These systems have proved invaluable for typing *E. coli* O157 isolates.

In the *E. coli* chromosome, *rec*A has an important role in SOS response and the recombination of damaged DNA during repair. The genes involved in this DNA repairing system are *rec*A, *uvr*A and *uvr*B and are regulated by *lex*A, which works as a repressor. The damaged DNA is reinitiated through pathways that involve these operons, which play a central role in recombination functions.[8]

Loss of *rec*A leads to improper processing of replication and causes generation of degradation-prone substrates.[9] Courcelle *et al*.[10] also suggested that RecA may be important for maintaining processive replication rather than to promote recombination.

Correspondence to: Professor John E. Moore
Northern Ireland Public Health Laboratory,
Department of Bacteriology, Belfast City Hospital, Belfast BT9 7AD,
Northern Ireland, UK
Email: jemoore@niphl.dnet.co.uk

The *rec*A gene has been used successfully for the epidemiological typing of other Gram-negative genera including *Burkholderia cepacia*[11] and *Erwinia* species.[12] However, to date, this locus has not been examined for *E. coli* O157 molecular epidemiology. Hence, it is the objective of this study to examine this locus by employing restriction fragment length polymorphism (RFLP), single-stranded conformational polymorphism (SSCP) and direct sequencing of the *rec*A PCR amplicon in a population of human clinical isolates obtained from faecal specimens of symptomatic patients.

Materials and methods

Bacterial isolates

E. coli O157:H7 isolates (n=80) were obtained from fresh faecal specimens throughout Northern Ireland and archived at the Northern Ireland Public Health Laboratory, Belfast

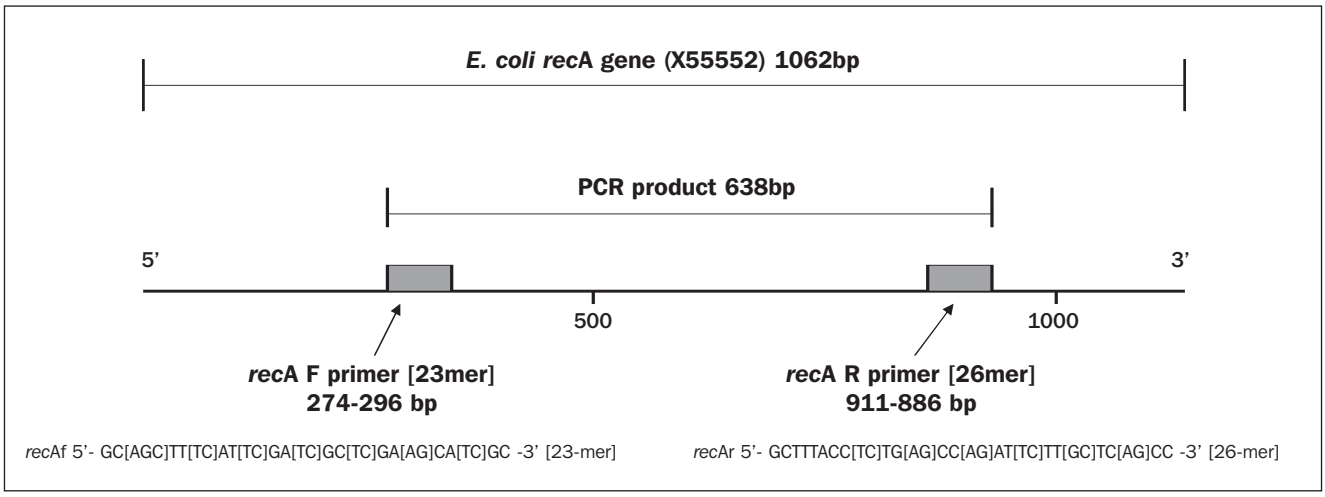

Fig. 1. Location of PCR primers on the *rec*A locus of *E. coli* (GenBank Accession No. X55552). [] = nucleotide degeneracy.

City Hospital, from 1997 to 1999. Twenty-eight isolates were recovered in 1997, 19 in 1998 and 33 in 1999. Isolates were stored in whole defibrinated horse blood (E&O Laboratories, Scotland) at –80°C until required. *E. coli* O157:H7 reference strain NCTC 12079 was used as a reference strain throughout this study.

Design of PCR assay

DNA sequence data of the *rec*A gene locus were obtained from GenBank (ww.ncbi.nlm.nih.gov/entrez) and conserved. Variable regions were identified subsequently by aligning 29 published *rec*A sequences (including *Acetobacter aceti, Acidothermus cellulyticus, Bifidobacterium adolescentis, B. anamalis, B. breve, Burkholderia cepacia, Campylobacter fetus, C. jejuni, Clostridium perfringens, Corynebacterium glutamicum, Enterobacter agglomerans, E. coli, Erwinia carotovara, Frankia alni, Gluconobacter oxydans, Proteus vulgaris, Provotella* spp., *Pseudomonas aeruginosa, P. fluorescens, P. putida, P. stutzeri, Salmonella typhimurium, Shigella flexerni, S. sonnei, Streptococcus parasanguis, Vibrio cholerae, Xanthomonas campestris, X. oryzae* and *Yersinia pestis*) by the Clustal alignment method employing the sequence alignment software package DNAStar (DNAStar, Wisconsin, USA). The novel degenerate primer pair, *rec*Af 5′- GC(AGC)TT(TC)AT(TC)GA(TC)GC(TC) GA(AG)CA(TC)GC-3′ (23-mer) and reverse – *rec*Ar 5′- GCTTTACC(TC)TG(AG)CC(AG)AT(TC)TT(GC)TC(AG)CC-3′ (26-mer), was designed to target conserved regions of the *rec*A gene for members of the Enterobacteriaceae The binding sites of the *rec*Af and *rec*Ar primers in *E. coli* (GenBank accession number X55552) relate to positions 274–296 and 911–886, respectively, yielding a fragment of 638 bp, as detailed in Figure 1, equating to approximately 60% coverage of the complete *rec*A locus in *E. coli*.

DNA extraction

All DNA isolation procedures were carried out in accordance with the DNA contamination management guidelines of Millar *et al*.[13] The High Pure PCR template preparation kit (Roche, UK) was used for extraction of the genomic DNA of *E. coli* isolates examined and was employed in accordance with the manufacturer's instructions. For health and safety reasons, *E. coli* O157:H7 was boiled for 10 min with TE buffer (100 µL) to heat kill all VTEC organisms prior to extraction.

recA PCR

Amplification reactions were set up in accordance with good molecular diagnostic procedures detailed in the guidelines of Millar *et al*.[13] All reaction mixes were set up in a PCR hood in a room separate from that used to extract DNA and from the amplification and post-PCR room, in order to minimise contamination. Initially, PCR amplification conditions were optimised by separately varying magnesium chloride concentration, annealing temperature, primer concentration and DNA template concentration.

Following optimisation, reaction mixes (100 µL) were set up as follows: 10 mmol/L Tris-HCl (pH 8.3), 50 mmol/L KCl, 2.5 mmol/L $MgCl_2$, 200 µmol/L (each) dATP, dCTP, dGTP and dTTP, 1.25 units *Thermus aquaticus* (*Taq*) DNA polymerase (Amplitaq, Perkin Elmer), 0.1 µmol/L each primer (*rec*Af and *rec*Ar) and 4 µL DNA template.

Following a 'hot start', the reaction mixtures were subjected to the following empirically optimised thermal cycling parameters in a Perkin Elmer 2400 thermocycler: initial 3 min of DNA denaturation at 94°C followed by 40 cycle of 96°C for 1 min, 58°C for 1 min and 72°C for 1 min, then 72°C for 10 min for final extension. Positive (*E. coli*) and multiple negative (water) amplification controls were included in every set of PCR reactions.

Following amplification, samples (10 µL) were removed from each reaction mixture and examined by electrophoresis (80V, 45 min) in gels composed of 2% (w/v) agarose (Gibco, UK) in TAE buffer (40 mmol/L Tris, 20 mmol/L acetic acid, 1 mmol/L EDTA [pH 8.3]), stained with ethidium bromide (5 µg/100 mL). Gels were visualised under ultraviolet (UV) illumination using a gel image analysis system (UVP Products, England) and all images were archived as digital graphic files (*.bmp).

Restriction fragment length polymorphism analysis

A restriction map was composed of the 638 bp of the *rec*A PCR amplicon sequence, employing DNAStar MapDraw software, and several enzymes were identified as having restriction sites within this amplicon. Among them, two enzymes which theoretically restrict the gene frequently (x7) were chosen for this study, including *Hha*I and *Hin*P1.

Separate DNA digestions were performed on each isolate genomic DNA at 37°C for 3 h and the digested DNA fragments were separated by electrophoresis. For the sample

preparation, 4 µL sample was mixed with 4 µL loading buffer (10 mmol/L Tris solution, 1 mmol/L EDTA solution, 1% [v/v] xylene cyanol solution and distilled water). For the DNA size marker, 2 µL 100-bp DNA ladder (Gibco BRL, UK) was mixed with 8 µL sample buffer.

The Multiphor II electrophoresis unit (Amersham Pharmacia LKB, UK) was connected to a thermostatic circulator and the temperature set to 15°C. For electrophoresis, the ExcelGel DNA analysis kit (Amersham Pharmacia Biotech AB, Uppsala, Sweden) was used. The cooling plate was moistened and a precast polyacrylamide gel (ExcelGel 48S, DNA, Amersham Pharmacia Biotech AB) was positioned in its centre. Cathodic and anodic buffer strips were applied to their respective sides of the gel. Prepared samples (7 mL) and DNA marker were applied in the well.

The electrodes were aligned with the centre of the buffer strips, and then the safety lid was placed in position. The power supply (Power Pac 3000, BioRad, USA) was connected to the electrodes and set to 600 V, 50 mA and 30 W. The electrophoresis was performed for 85 min. After electrophoresis, the gel was stained with a DNA silver staining kit (Amersham Pharmacia Biotech AB).

Single-stranded conformational polymorphism analysis

*rec*A PCR-RFLP was performed on each *E. coli* isolate, as described above, employing both restriction enzymes separately. After RFLP analysis was performed, the products were examined by SSCP analysis. The RFLP product (4 µL) was mixed with 4 µL denaturing solution (99% [v/v] formamide, 1% [v/v] xylene cyanol solution, 10 mg [0.1% v/w] bromophenol blue). The RFLP product was denatured at 96°C for 5 min and the PCR tube was cooled immediately to 4°C to prevent reannealing of the single-stranded DNA. Electrophoresis was performed as described above, except that SSCP samples were applied directly to wells without mixing with sample buffer.

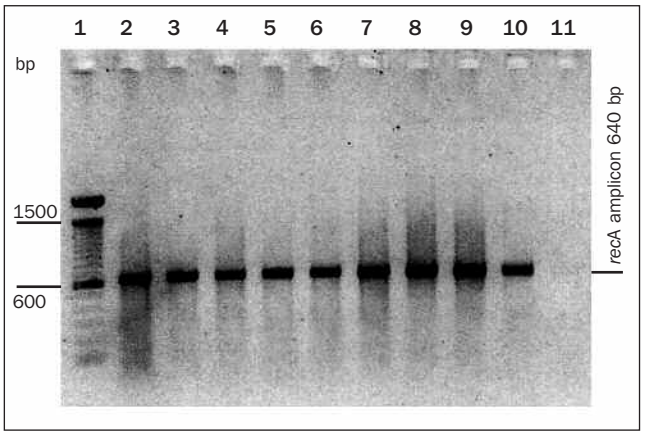

Fig. 2. PCR amplification of a 638 bp region of the *rec*A locus in nine wild-type clinical *E. coli* O157:H7 isolates. Lane 1: molecular weight marker (100 bp ladder, Gibco, Paisley, Scotland); lanes 2–10: clinical *E. coli* O157:H7; lane 11: negative control (molecular grade water, Biowhittaker, USA).

Results

Employing the *rec*A primer pair (*rec*Af/*rec*Ar), PCR was performed successfully on all 80 wild-type *E. coli* O157:H7 isolates, as well as on the reference strain (*E. coli* O157:H7 NCTC12079) and with 10 *E. coli* (non O157:H7), as well as with *Citrobacter freundii, Hafnia alvei, Shigella sonnei, Enterobacter cloacae* and *Providencia stuartii*, where a single band of the expected size (approximately 638 bp) was generated (Fig. 2).

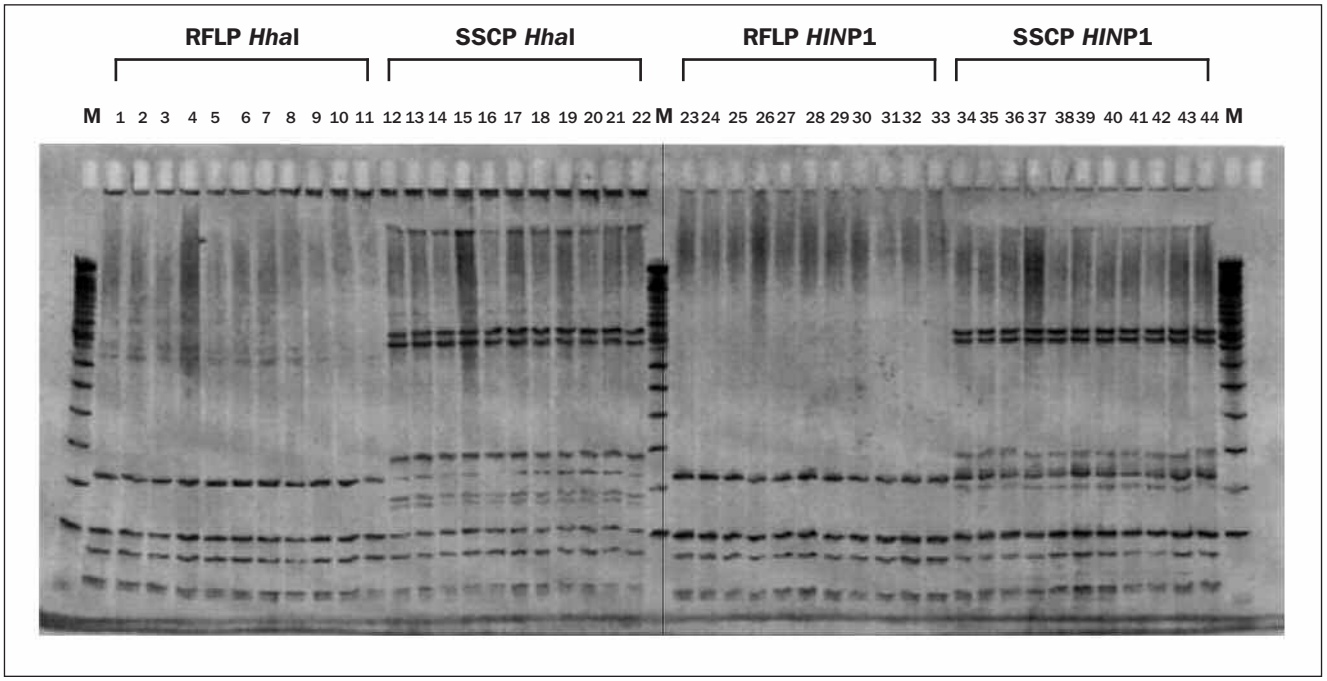

Fig. 3. Genotypic banding profiles of restriction fragment length polymorphism (RFLP) and single-stranded conformational polymorphism (SSCP) analyses of a partial *rec*A gene locus in 11 clinical isolates of *E. coli* O157:H7. M: molecular weight marker (100 bp ladder, Gibco, Paisley, Scotland); lanes 1–11: RFLP profile with *Hha*I; lanes 12–22: SSCP with *Hha*I profile of same isolates as in lanes 1–11; lanes 23–33: RFLP profile with *Hin*P1 of same isolates as in lanes 1–11; lanes 34–44: SSCP with *Hin*P1 profile of same isolates as in lanes 1–11.

To confirm the specificity of these primers as *rec*A, subsequent sequencing of the 638 bp *rec*A amplicon of a representative wild-type *E. coli* O157:H7 (E97/3N; VT2$^+$; eae$^+$; phage type 49) belonging to the common RFLP/SSCP genotype, which originated from a four-year-old male, was performed and confirmed the identity of the amplicon as *rec*A gene from *E. coli*. These sequence data have now been submitted to GenBank with the accession number EU344907, which was 100% identical with AY464233 (*E. coli* strain ECOR21 RecA protein [*rec*A] gene, partial cds).

Use of PCR-RFLP analysis demonstrated the presence of one *rec*A RFLP genotype for 79/80 *E. coli* O157:H7 isolates examined in 1997–1999 for *Hha*I restriction analysis, and a single profile for *Hin*P1. *Hha*I RFLP generated three fragments of approximately 60, 100 and 235 bp (Fig. 3). The clustering of isolates with the *Hin*P1 RFLP was identical to the *Hha*I RFLP (i.e., demonstrating the presence of a single genotype for the 79/80 *E. coli* O157:H7 isolates examined; Fig. 3). Although the two enzymes showed the same banding pattern by RFLP, *Hin*P1 I SSCP patterns were different from that of *Hha*I (Fig. 3).

Using *Hha*I and *Hin*P1 I enzymes, *E. coli* NCTC12079 displayed the same genotype as the wild-type *E. coli* O157 isolates for both RFLP and SSCP analysis. Analysis of other *E. coli* non-O157:H7, including *E. coli* O26, O86, O111, O125 and O128, demonstrated different banding patterns in both RFLP and SSCP analysis to the *E. coli* O157:H7, and their patterns were unique when compared to those of the *E. coli* O157 strains. Likewise, the five non-*E. coli* strains examined also demonstrated completely different banding patterns to *E. coli* O157 and *E. coli* non-O157 strains by both RFLP and SSCP analysis, whereby a common band (100 bp) was observed in *Citrobacter, Shigella* and *Enterobacter* by RFLP with *Hha*I and *Hin*P1.

Discussion

In the current study, novel PCR primers were designed to screen across approximately 60% of the complete *rec*A gene of *E. coli* (approximately 1062 bp) for the presence of potential polymorphisms between *E. coli* O157:H7 isolates, *E. coli* non-O157 isolates and non-*E. coli* isolates. In order to assess possible epidemiological diversity of *E. coli* O157 and non-O157 isolates, as well as non-*E. coli* isolates, PCR-RFLP and PCR-RFLP-SSCP techniques were employed using the *rec*A gene locus.

In order to accomplish this objective, a novel PCR assay had to be designed whereby a novel *rec*A primer pair was designed to amplify a wide range of the Enterobacteriaceae family. From available alignment data, several degeneracies had to be incorporated into the design of the forward and the reverse primers (Fig. 1), which subsequently allowed the successful amplification of other members of the Enterobacteriaceae family, as well as *E. coli* O157:H7.

Overall, RFLP and SSCP analysis demonstrated the presence of a single genotype in 79/80 of the wild-type *E. coli* O157:H7 isolates examined and different profiles for non O157 *E. coli* and other non-*E. coli* organisms. There was a single wild-type *E. coli* O157:H7 isolate (E97/24E) that did not conform to the RFLP and SSCP profiles obtained from all the other *E. coli* O157:H7 clinical isolates.

Further analysis revealed that this isolate originated from an eight-month-old baby boy, and the isolate was successfully speciated as *E. coli* using the API20E identification scheme (5144572) and was negative by PCR analysis for the presence of *VT*1 and *VT*2 gene loci, but was positive for the *eae* gene locus.[14] In addition, this isolate was non-typable by phage-typing, giving the 'reacts but does not conform' (RDNC) subtype, as well as a unique PFGE genotype not observed in any other local *E. coli* O157 organism. Thus, given the atypical nature of this strain, additional phenotypic tests are warranted to confirm its status as a verocytotoxigenic O157:H7 organism.

Various molecular approaches may be useful in detecting polymorphisms within hypervariable regions of a target gene such as the *rec*A locus. Such approaches include RFLP, SSCP and sequence typing. Previously, SSCP has been described as a useful method for fine and sensitive discrimination of closely related taxa of a variety of bacterial and viral gene loci.[15,16] In addition, the *rec*A gene has been employed previously to help differentiate and characterise several bacterial genera, including *Acinetobacter* and *Borrelia*.

In the current study, novel PCR primers were designed to screen across approximately 60% of the complete *rec*A gene of *E. coli* (approximately 1062 bp). Most recently, use of the *rec*A locus has found application in several multilocus sequence typing (MLST) methods, including with *Listeria monocytogenes*,[17] *Yersinia* species,[18] *Mycoplasma hyopneumoniae*[19] and *Porphyromonas gingivalis*,[20] where the range of the number of polymorphic sites has varied (88, 92, 10 and 2 polymorphic sites, respectively).

In conclusion, this study employed a novel PCR assay to generate *rec*A amplicons of approximately 638 bp in length. On subsequent RFLP and SSCP analysis they were shown locally to be highly clonal in a population of 79 human clinical *E. coli* O157 organisms isolated in Northern Ireland during a three-year period (1997–1999) over which verocytotoxigenic infection in this region increased markedly. □

The authors wish to thank all clinical microbiology laboratories in Northern Ireland for the forwarding of all their E. coli O157:H7 isolates for molecular typing. JEM, MASMcM and DAMcD are funded by a Research & Development Office Grant (RRG 9.9).

References

1 Rangel JM, Sparling PH, Crowe C, Griffin PM, Swerdlow DL. Epidemiology of *Escherichia coli* O157:H7 outbreaks, United States, 1982–2002. *Emerg Infect Dis* 2005; **11**: 603–9.

2 Michino H, Araki K, Minami S *et al*. Massive outbreak of *Escherichia coli* O157:H7 infection in schoolchildren in Sakai City, Japan, associated with consumption of white radish sprouts. *Am J Epidemiol* 1999; **150**: 787–96.

3 Cowden JM, Ahmed S, Donaghy M, Riley A. Epidemiological investigation of the central Scotland outbreak of *Escherichia coli* O157 infection, November to December 1996. *Epidemiol Infect* 2001; **126**: 335–41.

4 Rogerie F, Marecat A, Gambade S, Dupond F, Beaubois P, Lange M. Characterization of Shiga toxin-producing *E. coli* and O157 serotype *E. coli* isolated in France from healthy domestic cattle. *Int J Food Microbiol* 2001; **63**: 217–23.

5 Aleksic S, Karch H, Bockemuhl J. Biotyping scheme for Shiga-like (Vero) toxin-producing *Escherichia coli* O157 and a list of

serological cross-reactions between O157 and other Gram-negative bacteria. *Zentralbl Bakteriol* 1992; **276**: 221–30.

6 Chinen I, Tanaro JD, Miliwebsky E *et al*. Isolation and characterization of *Escherichia coli* O157:H7 from retail meats in Argentina. *J Food Prot* 2001; **64**: 1346–51.

7 Lin CK, Lin JC. Development of PCR primers based on a fragment from randomly amplified polymorphic DNA for the detection of *Escherichia coli* O157:H7/NM. *Mol Cell Probes* 2007; **21**: 182–9.

8 Paul R, Dalibart R, Lemoine S, Lestienne P. Expression of *E. coli* RecA targeted to mitochondria of human cells. *Mutat Res* 2001; **486**: 11–9.

9 Yassien MA, Ewis HE, Lu CD, Abdelal AT. Molecular cloning and characterization of the *Salmonella enterica* serovar paratyphi B *rma* gene, which confers multiple drug resistance in *Escherichia coli*. *Antimicrob Agents Chemother* 2002; **46**: 360–6.

10 Courcelle J, Ganesan AK, Hanawalt PC. Therefore, what are recombination proteins there for? *Bioessays* 2001; **23**: 463–70.

11 Crowley D, Daly M, Lucey B *et al*. Molecular epidemiology of cystic fibrosis-linked *Burkholderia cepacia* complex isolates from three national referral centres in Ireland. *J Appl Microbiol* 2002; **92**: 992–1004.

12 Waleron M, Waleron K, Podhajska AJ, Lojkowska E. Genotyping of bacteria belonging to the former *Erwinia* genus by PCR-RFLP analysis of a *rec*A gene fragment. *Microbiology* 2002; **148**: 583–95.

13 Millar BC, Xu J, Moore JE. Risk assessment models and contamination management: implications for broad-range ribosomal DNA PCR as a diagnostic tool in medical bacteriology. *J Clin Microbiol* 2002; **40**: 1575–80.

14 Crothers JL, Moore JE, Millar BC, Watabe M, Rooney PJ. Determination of verocytotoxin and *eae* gene loci by multiplex PCR in *Escherichia coli* O157:H7 isolated from human faeces in Northern Ireland: a four-year study of trends, 1997–2000. *Br J Biomed Sci* 2004; **61**: 1–7.

15 Jawad A, Snelling AM, Heritage J, Hawkey PM. Comparison of ARDRA and *rec*A-RFLP analysis for genomic species identification of *Acinetobacter* spp. *FEMS Microbiol Lett* 1998; **165**: 357–62.

16 Kerr JR, Curran MD, Moore JE *et al*. Genetic diversity in the non-structural gene of parvovirus B19 detected by single-stranded conformational polymorphism assay (SSCP) and partial nucleotide sequencing. *J Virol Methods* 1995; **53**: 213–22.

17 Jiang, L., Chen, J., Xu, J *et al*. Virulence characterization and genotypic analyses of *Listeria monocytogenes* isolates from food and processing environments in eastern China. *Int J Food Microbiol* 2008; **121** (1): 53–9.

18 Kotetishvili M, Kreger A, Wauters G, Morris JG, Sulakvelidze A, Stine OC. Multilocus sequence typing for studying genetic relationships among *Yersinia* species. *J Clin Microbiol* 2005; **43**: 2674–8.

19 Mayor D, Jores J, Korczak BM, Kuhnert P. Multilocus sequence typing (MLST) of *Mycoplasma hyopneumoniae*: a diverse pathogen with limited clonality. *Vet Microbiol* 2008; **127** (1–2): 63–72.

20 Enersen M, Olsen I, van Winkelhoff AJ, Caugant DA. Multilocus sequence typing of *Porphyromonas gingivalis* strains from different geographic origins. *J Clin Microbiol* 2006; **44**: 35–41.